New Phytologist

Research

Enoyl-CoA hydratase genes underlie natural variation of lactones in cultivated and wild strawberries

Mark Porter[1], Xuebo Song[2], Jin-Pyo An[2], Maria G. Blanco[3], Zhen Fan[4], Yu Wang[2], Timo Hytönen[5], Carmen Martín-Pizarro[3], Vance M. Whitaker[1], David Posé[3] and Seonghee Lee[1]

[1]Horticultural Sciences Department, University of Florida, IFAS Gulf Coast Research and Education Center, 14625 CR 672, Wimauma, FL 33598, USA; [2]Food Science Department, University of Florida IFAS Citrus Research and Education Center, 700 Experiment Station Rd, Lake Alfred, FL 33850, USA; [3]Departamento de Biología Molecular y Bioquímica, Facultad de Ciencias, UMA, Instituto de Hortofruticultura Subtropical y Mediterránea La Mayora (IHSM La Mayora), Universidad de Málaga – Consejo Superior de Investigaciones Científicas (UMA-CSIC), Málaga, 29010, Spain; [4]College of Horticulture Sciences, Zhejiang A&F University, Hangzhou, Zhejiang, 311300, China; [5]Department of Agricultural Sciences, Viikki Plant Science Centre, University of Helsinki, 00014, Helsinki, Finland

Authors for correspondence:
Seonghee Lee
Email: seonghee105@ufl.edu

Vance M. Whitaker
Email: vwhitaker@ufl.edu

David Posé
Email: dpose@uma.es

Received: *17 September 2025*
Accepted: *16 February 2026*

New Phytologist (2026)
doi: 10.1111/nph.71070

Key words: eQTL, fruit breeding, fruit flavor, GWAS, haplotype-phased genome assembly, lactone biosynthesis.

Summary

- Strawberry flavor is a key factor in consumer acceptance and is highly influenced by volatile aroma compounds. Multiple consumer and trained panel sensory studies indicate that γ-lactones are among the most important aroma compounds in strawberry, imparting sweet, fruity aromas.
- We quantified γ-lactones in *Fragaria* × *ananassa* and *Fragaria vesca* to explore the genetic basis for lactone production in the two species using genome wide association study (GWAS). We developed a haplotype-phased genome assembly for a high lactone cultivar to explore previously published expression quantitative trait locus and fruit ripening RNA sequencing data to identify candidate genes in *Fragaria* × *ananassa*. We leveraged knowledge from *Fragaria* × *ananassa* to identify candidate genes in *F. vesca* and validated one gene from each strawberry species using fruit transient gene expression assays.
- We found that a novel Enoyl CoA hydratase/isomerase gene in *Fragaria* × *ananassa*, *FaECH*, increases concentration of γ-lactones while a similar but nonhomologous gene, *FvECH*-like, increases γ-lactones in *F. vesca*. Additionally, the major gene shown to control γ-decalactone production in *Fragaria* × *ananassa*, *FaFAD1*, has a homolog in *F. vesca*, *FvFAD1*, that underlies a strong GWAS.
- These results validate key γ-lactone biosynthesis steps in both strawberry species and provide tools to breed more flavorful strawberries.

Introduction

Both cultivated (*Fragaria* × *ananassa*, $2n = 8x = 56$) and woodland strawberries (*Fragaria vesca*, $2n = 2x = 14$) produce hundreds of volatile aroma compounds, of which fewer than 30 are likely to contribute to flavor (Ulrich *et al.*, 2018; Porter *et al.*, 2023; Urrutia *et al.*, 2023). Understanding the key aroma compounds in strawberry fruit has been an evolving field of study for over 30 yr (Ulrich *et al.*, 2018). From quantification of volatile compounds and calculation of theoretical odor activity values, to combining large sensory and chemical datasets, our knowledge of important aroma compounds is becoming more complete (Schwieterman *et al.*, 2014; Ulrich & Olbricht, 2016; Fan *et al.*, 2021). Furanones, lactones, esters and aldehydes have the largest impact on sensory attributes, mostly contributing positively. γ-decalactone and γ-dodecalactone are the most abundant and perceptible lactones in strawberry fruit that impart peach-like, fruity aromas, which positively affect sweetness, liking and flavor perception in sensory panels (Ulrich & Olbricht, 2016; Fan *et al.*, 2021). γ-dodecalactone and γ-decalactone were correlated with aromatic ($r = 0.69$ for γ-dodecalactone vs $r = 0.41$ for γ-decalactone) and acceptance ($r = 0.61$ for γ-dodecalactone vs $r = 0.47$ for γ-decalactone) sensory attributes in a European consumer panel study (Ulrich & Olbricht, 2016). γ-dodecalactone was also identified as an important predictor of sweetness and liking in a different consumer panel study, whereas γ-decalactone was only an important predictor of sweetness (Fan *et al.*, 2021). Trained sensory panels identified γ-dodecalactone and γ-decalactone as important predictors of sweetness and strawberry flavor, respectively, in two out of the six years of sensory evaluation (Fan *et al.*, 2021). Therefore, identifying and characterizing genes involved in lactone biosynthesis is of great interest to facilitate the breeding of more flavorful strawberries.

A major gene controlling γ-decalactone biosynthesis in cultivated strawberry, *FaFAD1*, was mapped on octoploid strawberry chromosome 3B, cloned and validated in transient fruit assays

(Chambers *et al.*, 2014; Sánchez-Sevilla *et al.*, 2014; Oh *et al.*, 2021). *FaFAD1* is responsible for increasing γ-decalactone to perceptible levels in a dosage-dependent fashion, and codominant high-resolution melting and PCR Allele Competitive Extension (PACE) markers were developed as breeding tools (Oh *et al.*, 2021). Aside from *FaFAD1*, only a few studies have investigated the genetics of lactone production in strawberries. A QTL mapping study in cultivated strawberry found a somewhat stable QTL for γ-lactones on chromosome 6D (Rey-Serra *et al.*, 2022). Another study investigated the transcriptomic basis of lactone biosynthesis in *Fragaria nilgerrensis* and derived populations showed that δ-lactones and not γ-lactones were important for discriminating genotypes with high and low peach-like aroma (Wang *et al.*, 2022). The authors report 18 genes correlated with δ-lactone concentration, including four shared structural genes, a fatty acid desaturase, an allene oxide synthase and a cytochrome P450 gene. Additionally, two *NAC* and one *MYB* family genes were correlated with δ-lactones. γ-decalactone was found to have a significant genetic component driving its variation in a QTL mapping study of *F. vesca*, but no major QTLs were discovered (Urrutia *et al.*, 2017).

Lactones are present in a wide range of species, from the yeast *Yarrowia lipolytica* to the popular fruit with which they are most associated: peach (*Prunus persica*), which shares the Rosaceae family with strawberry. Groundwork for elucidation of the lactone pathway in fruits comes from Schöttler and Boland, who demonstrated the formation of an artificial undecanolactone by feeding strawberry and nectarine with labeled epoxy acid of *cis* 9,10-([^{18}O ~ Epoxy)heptadecanoic acid (Schöttler & Boland, 1996). The major components of the lactone pathway involve desaturation of fatty acids, epoxidation/hydration of double bonds resulting in hydroxy fatty acids, β-oxidation and finally cyclization. The lactone pathway has been well- studied in peach with three genes, a fatty acid desaturase (*PpFAD2-1*), acyl-CoA oxidase (*PpACX1*) and an alcohol acyltransferase (*PpAAT1*), functionally validated (Li *et al.*, 2023). Notably, the homolog of *FaFAD1* in peach, *PpFAD2-1*, was functionally characterized with ω-6 oleate desaturase activity, increasing linoleic acid (18:2) concentration *in vitro* and establishing linoleic acid as an important precursor, especially for γ-decalactone (Sánchez *et al.*, 2013; Peng *et al.*, 2023). Additionally, transient overexpression of *PpFAD2-1* in peach fruit increased γ-decalactone, γ-dodecalactone, γ-octalactone and γ-hexalactone concentrations (Peng *et al.*, 2023). *PpACX1* is a key enzyme that activates fatty acyl-CoA precursors for β-oxidation (Zhang *et al.*, 2017), and *PpAAT1* catalyzes the last step in the lactone pathway with esterification of 4-hydroxyacyl-CoAs in the case of γ-lactones (Peng *et al.*, 2020).

Here, we use a large breeding population of 406 individuals to investigate genetic variation contributing to lactone concentration in cultivated strawberry fruit. Use of previous eQTL data with functional annotations revealed a candidate gene underlying a genome wide association study (GWAS) peak encoding a gene with an Enoyl-CoA hydratase/isomerase domain, *FaECH*. Fruit transient RNAi and overexpression experiments demonstrated the ability of *FaECH* to modulate γ-dodecalactone, γ-decalactone and γ-octalactone. Additionally, we use a previously characterized European diversity collection of *F. vesca* to identify genetic mechanisms of lactone production in this diploid species, revealing a strong effect of the *FaFAD1* homolog, *FvFAD1*. Interestingly, a novel association on *F. vesca* chromosome 7 was also found to underlie γ-decalactone production and localized to a gene with an Enoyl-CoA hydratase/isomerase domain, *FvECH*-like, distinct from *FaECH* in *F.* × *ananassa*. Integrating knowledge of lactone biosynthesis from peach, we outline key genes in the strawberry lactone biosynthesis pathway using genetic associations in this study and suggest other strawberry genes that may also contribute to lactone biosynthesis.

Materials and Methods

Plant materials

Fragaria × *ananassa* Duchesne ex Rozier ($2n = 8x = 56$) plants were grown at the Gulf Coast Research and Education Center (Wimauma, FL, USA) during the 2021–2022 strawberry season according to commercial standards. A total of 385 advanced selections and 21 named varieties were chosen to capture the diversity of the University of Florida breeding program with minimal population structure (Supporting Information Table S1; Fig. S1). Plants were grown in 10-plant plots, and three to four fully ripe (at least 90% red) fruit were harvested from each plot on the morning of 11 January 2022 and immediately frozen in liquid nitrogen before storage in −20°C.

Fragaria vesca L. ($2n = 2x = 14$) plant materials are detailed in Urrutia *et al.* (2023). Briefly, three clones of 125 and 170 natural accessions were grown under polycarbonate during the summer of 2016 and 2017, respectively, and at least five fully ripe berries were harvested from each genotype. Harvested fruits were immediately frozen in liquid nitrogen before being ground to a fine powder and stored at −80°C.

Quantification of γ-lactones in *Fragaria* × *ananassa*

For the GWAS population, γ-lactones were extracted from fruit tissue using a liquid–liquid extraction method with minor modifications to a method previously described (Song *et al.*, 2023). Five grams of homogenized strawberry was transferred to a 15-ml glass tube and spiked with 20 µl of [$^{8}H_2$]-naphthalene (internal standard, 12.5 mg l^{-1}). The mixture was vortexed for 15 s, followed by the addition of 5 ml dichloromethane/hexane (1 : 1, v/v) and vortexing for another 30 s. Samples were subsequently agitated on an oscillator (speed setting 2) for 3 min and sonicated for 15 min at room temperature. Phase separation was achieved by centrifugation at 1860 g for 15 min at 4°C. The upper organic phase was collected into clean glass tubes, and the combined extracts were evaporated under a gentle nitrogen stream to a final volume of *c.* 0.5 ml. Quantification was based on standard calibration curves with [$^{8}H_2$]-naphthalene as the internal standard. To obtain a blank strawberry matrix, commercial strawberries (variety unknown) were purchased from Walmart (Grand Blend). 100 grams of homogenized strawberry was extracted

three times with dichloromethane (200 ml each) under continuous stirring for 2 h per extraction. After centrifugation, the solid residue was collected and dried in an oven. The dried strawberry material was then reconstituted with water at a weight ratio of 1 : 9 (solid : water). This reconstructed matrix was subjected to the same extraction procedure as described previously to confirm that no target lactones could be detected and was subsequently used as the blank matrix in this study. The known amounts of lactone compounds in different concentrations and the 20 µl of [$^{8}H_2$]-naphthalene (12.5 mg l^{-1}) were added to the blank matrix sample. After equilibration, the samples with [$^{8}H_2$]-naphthalene and lactone compounds were processed the same as liquid–liquid extraction (LLE). The standard curves were prepared by plotting the ratio of the standard compounds and the internal standard against their concentration ratio. The prepared strawberry extracts were stored at −80°C until further analysis by GC-MS/MS. For analysis, 1.0 µl of the extracted sample was introduced into the GC-MS/MS (Song *et al.*, 2023).

Quantification of γ-lactones in *Fragaria vesca*

Volatile analysis of strawberry fruit was conducted according to Urrutia *et al.* (2023). Briefly, lactones were sampled by headspace solid-phase microextraction (SPME) using a 65 µm polydimethylsiloxane/divinylbenzene (PDMS/DVB) fiber (Supelco, Bellefonte PA, USA). GC-MS was performed in a 6890 N gas chromatograph coupled to a 5975B mass spectrometer (Agilent Technologies, Santa Clara CA, USA). Compounds were monitored over the mass/charge ratio ($mz - 1$) range of 35–250. Chromatograms and mass spectra were analyzed using the ENHANCED CHEMSTATION software (Agilent Technologies). Volatile organic compounds (VOCs) were unambiguously identified by comparison of both retention time and mass spectra to those of commercial standards (Sigma-Aldrich) run under the same conditions. For quantification, a specific ion was selected for integration of the area of each of the identified compounds. Areas were normalized by comparison with the peak area of the same compound in the flanking reference samples. The reference sample consisted of a homogeneous mix of all analyzed samples and was injected regularly every five to six samples in order to correct for variations in sensitivity and fiber aging. Volatile analysis for transient assay fruit was carried out using SPME following the same methods as in Pott *et al.* (2021).

Genome assembly

Etiolated leaf tissue was prepared from in-field plants using black plastic pots to exclude light for 2 wk before etiolated leaves were frozen and submitted to DNA Link (Orlando, FL, USA) for genomic DNA extraction, library preparation and sequencing. The prepared library was sequenced using two 8 M SMRT cells using v.2 chemistry and v.2.1 polymerase. The total yield of HIFI data was 48.5 Gb with an average read length of 11.6 kb. Medallion™ 'FL 16.30–128' (hereafter referred to as Medallion for simplicity), and its parents FL 12.90–53 and FL 13.27–142 were sequenced with Illumina NovaSeq 150 bp paired-end. A total of 13, 11.9 and 12.3 Gb of paired-end data were generated for Medallion, FL 12.90–53 and FL 13.27–142, respectively.

A trio-binning approach was carried out using HIFIASM v.0.16.1 with YAK v.0.1 (Cheng *et al.*, 2021) to assemble both haplotypes of Medallion using parental Kmers. After assembly, both haplotype assemblies were corrected using RAGTAG v.2.0.1 based on the 'Royal Royce' genome (FaRR1) (Hardigan *et al.*, 2021) and supplied with Medallion Illumina reads to limit incorrect breakpoints (Alonge *et al.*, 2022). After correction, both haplotypes were scaffolded to the primary 'Royal Royce' assembly using RAGTAG v.2.0.1 to construct pseudochromosomes (Alonge *et al.*, 2022).

Before annotation, a custom repeat library was created using REPEATMODELER v.2.0, and each haplotype genome was masked using REPEATMASKER v.4.1.1. The Braker3 pipeline was followed using BRAKER v.3.0.8 with ISOseq and proteins as evidence. First, all available *Fragaria* × *ananassa* Isoseq datasets (Table S2) were downloaded from the NCBI SRA and mapped to each Medallion haplotype genome using MINIMAP v.2.24 and concatenated. The Viridiplantae ORTHODB v.11 protein set was downloaded (https://bioinf.uni-greifswald.de/bioinf/partitioned_odb11/), and Braker was run using isoseq and protein evidence to annotate both haplotype genomes independently. Details of assembly and annotation along with quality evaluation are outlined in Notes S1.

Additionally, for eQTL mapping and GWAS, positions of single nucleotide polymorphisms (SNPs) in the Axiom™ Strawberry FanaSNP 50 k array were identified in the Medallion haplotype 1 genome by BLAST search of SNP probe flanking sequences (Camacho *et al.*, 2009). Probe hits mapping to a single chromosome were retained while hits for probes mapping to multiple chromosomes were assigned to the chromosome in 'Royal Royce' or when there was no match; the hit with the highest bit score was retained (Dataset S1).

Nucleotide and protein sequence alignments

All nucleotide alignments and BLAST searches were carried out using GENEIOUS PRIME 2024.0.7 (https://www.geneious.com). Nucleotide alignment was generated with the CLUSTAL OMEGA 1.2.2 function while BLAST was done using the BLAST function.

Genome-wide association studies

For cultivated strawberry genotypes, DNA was extracted from leaf tissue before genotyping with the Axiom™ Strawberry FanaSNP 50 k Genotyping Array (Hardigan *et al.*, 2019). Missing markers were imputed using BEAGLE v. 5.2 (Browning *et al.*, 2018). The relatedness matrix was calculated using the R package RMVP v.1.1.1 before association testing for each phenotype using the mixed linear model (MLM) method with the kinship matrix, top 3 principal components and 42 543 SNPs with MAF > 0.01 (Yin *et al.*, 2021). Manhattan plots were created using the R package CMPLOT (Yin *et al.*, 2021) using physical position of markers in the Medallion haplotype 1 genome.

For *F. vesca* natural accessions, genotyping was carried out through Illumina whole genome sequencing (NextSeq 500; 150 + 150 paired end reads). DNA extraction, whole genome sequencing, read mapping and variant calling are detailed in Toivainen *et al.* (2026). RMVP v.1.1.1 was used to conduct association testing using 2720 726 markers and 121 individuals for the 2016 data (H16) and 168 for the 2017 data (H17). Lactone peak area ratios were normalized by $\log_2$ transformation and rescaled using the formula: function(x) = (x–min(x))/(max(x)–min(x)) as in Urrutia *et al.* (2023). Statistical analysis was carried out using the R software basic functions unless otherwise is specified. MLM and FarmCPU models were used to identify significant SNPs using the Bonferroni- corrected threshold of P = 1.84e-08.

Cis eQTL mapping

Raw RNA sequencing reads were retrieved from the NCBI SRA database (Sayers *et al.*, 2022), and a total of 170 red fruit samples consisting of 83 California and 87 Florida genotypes were retained for eQTL mapping (Fan *et al.*, 2022) (Table S3). Adapters and low- quality reads were removed with TRIMMOMATIC v.0.39 before mapping to Medallion haplotype 1 with STAR v.2.7.11b (Dobin *et al.*, 2013). Only single-mapping reads were used to count genomic features using FEATURECOUNTS in the SUBREAD v.2.0.6 package (Liao *et al.*, 2014). DESEQ2 was used to normalize raw gene counts by the median of ratios method and filtering all genes with a sum of normalized counts < 10 (Love *et al.*, 2014). *Cis* eQTL mapping was conducted using QTLTOOLS v.1.3.1 by first conducting principal component analysis on the matrix of 53 909 genes then mapping eQTLs for each gene within a 1000 000-bp window using 43 394 SNPs and 6 PCs as covariates (Delaneau *et al.*, 2017).

Gene expression during fruit ripening

The NCBI SRA was searched for all *Fragaria* × *ananassa* short-read transcriptomic data, and the SRA run table was downloaded. Manual filtering for datasets generated on fruit samples at different ripening stages yielded 89 different 150-bp paired-end RNA-seq datasets for 10 different cultivars (Table S4). Raw RNA sequencing reads were downloaded and processed in the same way as described previously for the eQTL, except no genes were filtered after DESEQ2 normalization. Although there was heterogeneity of ripening stages across samples, four ripening stages were defined, and green, white, turning and red samples were assigned to each category depending on their developmental stage from the SRA metadata (Table S4). Ripening upregulation was determined by using the contrast function in the DESEQ2 between the 'green' and 'red' categories for each gene.

Transient fruit gene expression experiments

Transient expression assays in detached *Fragaria* × *ananassa* fruits were conducted similar to Pi *et al.* (2019) with modifications (Jang *et al.*, 2024). The full-length coding sequence (CDS) of *FaECH* was determined by finding the longest open reading frame from 'Florida Brilliance' IsoSeq reads that covered the entire gene model. 350-bp partial CDS was optimized for RNAi by pssRNAit (Ahmed *et al.*, 2020). Full-length CDS and partial CDS were flanked by attB1 and attB2 for Gateway cloning and synthesized by Twist Biosciences, San Francisco, CA (Table S5). Fragments were cloned into gateway destination vectors pMDC32 for overexpression and pK7GWIWG2(II) for RNAi, respectively. *FaECH*-OX, *FaECH*-RNAi, and both empty pMDC32 and pK7GWIWG2(II) vectors were transformed into *Agrobacterium tumefaciens* strain EHA105. White fruits were harvested from field- grown plants of Medallion for RNAi and Ember™ 'FL 20.80-4' (a lower lactone producer) for overexpression. Agrobacterium suspension was injected into fruits using a 5-ml syringe until full saturation. At 5 d post injection, 2–3 fruits per treatment were pooled as technical replicates and were flash- frozen in liquid nitrogen and ground to powder before storage at −80°C. RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich), and quantitative polymerase chain reaction (qPCR) was carried out on cDNA according to Damgaard & Treebak (2022) using the primers for *FaECH* and *GAPDH2* in Table S5. Volatile analysis was performed using SPME (Notes S2).

Overexpression of *FvECH*-like in *F. vesca* was performed following the same procedure used for the cultivated strawberry. The full-length CDS of *FvECH*-like was cloned into pK7WG2 vector and introduced into *A. tumefaciens* strain GV3101. Green fruits from *F. vesca* 'Reine des Vallés' were injected with Agrobacterium suspension and allowed to ripen on the plant. Red fruits (5–7 d after infection) were harvested, and immediately frozen in liquid nitrogen before storage at −80°C. qPCR was carried out on RNA extracted from 10 fruit samples using the comparative threshold cycle method ($\Delta\Delta C_T$) with primers specific for *FvECH*-like (Table S5). For volatile analysis, five groups of fruits were paired for the control and overexpression treatments (Table S6). Volatile analysis was carried out similar to the methods mentioned earlier but with some modifications (Notes S2).

Results

Genome Assembly for High-Lactone Cultivar Medallion™ 'FL 16.30-128'

Medallion is notable for its exceptional flavor resulting from high soluble solids and abundant lactones (Dalid *et al.*, 2025). Trained sensory panel scores for 'strawberry flavor' and 'sweetness' are higher for Medallion than other commercial cultivars such as 'Florida Brilliance' over multiple years (Fig. 1a). Medallion carries two copies of *FaFAD1* on chromosome 3B (Fig. 1b) and is the first cultivar release from the University of Florida breeding program with perceptible lactone aroma resulting from the use DNA markers to select for seedlings with two doses of the functional *FaFAD1* allele (Oh *et al.*, 2021).

After long-read sequencing and assembly, both Medallion haplotypes were assembled with an N50 of 10.7 and 16.7 Mb for Haplotype 1 and 2 and scaffolded length of 717.2 and 748.6 Mb

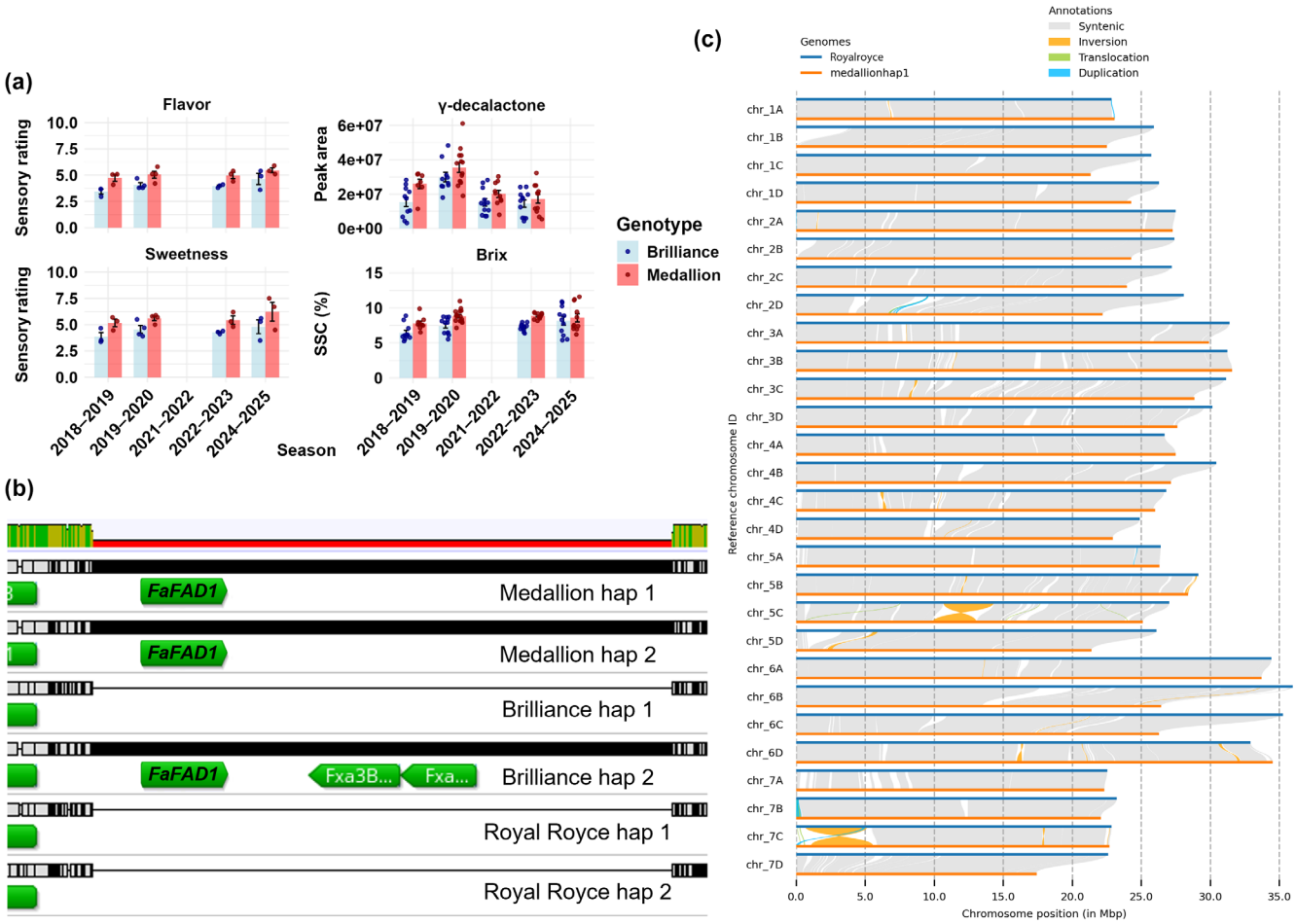

Fig. 1 Medallion strawberry (*Fragaria* × *ananassa*) fruit quality comparison and genome alignments. (a) Bar charts on the left side display trained panelist sensory ratings of 'strawberry flavor' and 'sweetness' for Medallion and 'Florida Brilliance' fruits across multiple seasons. Bar charts on the right side show γ-decalactone concentration and soluble solids content. The error bars denote 1SE. Data were taken from trained descriptive analysis panels conducted in Florida (Fan *et al*., 2021). (b) Pairwise alignment of Medallion, 'Florida Brilliance' and 'Royal Royce' at the *FaFAD1* locus. Light gray indicates identical nucleotide bases, black boxes indicate different nucleotide bases, and black lines indicate an absence of nucleotide bases. (c) Synteny plot generated using SyRI (Goel *et al*., 2019) between Medallion haplotype 1 and the 'Royal Royce' primary assembly.

as well as LTR Assembly Index (LAI) peroxisomal targeting signal-1 scores of 16.45 and 16.13 (Table 1). Additionally, 99.0% and 99.2% of total protein CDSs in the embryophyta_odb10 database in BUSCO were complete. A total of 69 825 and 73 379 genes were annotated in each haplotype with 98.1% complete BUSCO sequences in both haplotype gene sets. Furthermore, there was a high degree of synteny between the 'Royal Royce' primary assembly and Haplotype 1 of Medallion (Fig. 1c).

Lactone production in *Fragaria* × *ananassa* is dependent on *FaFAD1* and a novel locus on chromosome 7A

We conducted volatile profiling and genotyping of 406 breeding lines and named varieties representing the diversity of the University of Florida strawberry breeding program in order to identify SNPs and candidate genes associated with lactone production. Genome-wide association showed a single significant peak on

Table 1 Genome assembly statistics for both haplotypes of the Medallion genome.

Assembly statistic	Medallion Haplotype 1 assembly	Medallion Haplotype 2 assembly
Contig size (Mb)	852.3	808.8
Number of contigs	6067	1889
Scaffolded size (Mb)	717.2	748.6
Number of scaffolded contigs	182	115
Largest contig (Mb)	34.5	32.6
N50 (Mb)	10.7	16.7
Completeness (%)[a]	97.1	99.2
BUSCO (%)	99	99.2
BUSCO frag (%)	0.1	0.1
QV	58	61

[a]Completeness refers to hapmer recovery in each assembly calculated in Merqury.

chromosome 3B significantly associated with γ-decalactone and γ-octalactone abundance localizing to the previously characterized *FaFAD1* gene responsible for dramatic increases in γ-decalactone production in strawberry (Fig. 2). Due to high homology among subgenomes, some significant markers are misplaced on homologous chromosome 3A (Fig. 2). There was no significant GWAS between the *FaFAD1* region and γ-dodecalactone production; however, a significant association was observed for γ-dodecalactone on chromosome 7A (Fig. 2). To understand the effect of each locus on the production of lactones, each combination of alleles at SNPs AX184908435 (3B) and AX123363242 (7A) was plotted against the concentrations of the lactones (Fig. 3) and one- and two-way Analysis of Variance (ANOVA) conducted for both loci and all lactones (Table S7).

A large dosage effect is observed on γ-decalactone and γ-octalactone concentration for the 3B locus consistent with the known effect of *FaFAD1* on γ-decalactone (Fig. 3a). Group means for zero and two doses of the alternative allele were significantly different with mean γ-decalactone rising from 0.3310 to 4.1883 mg kg^{-1} and mean γ-octalactone from 0.0085 to 0.0300 mg kg^{-1}, a 12-fold and 3.5-fold increase, respectively (Table S8). The 3B locus shows a smaller but significant effect on γ-dodecalactone production in single marker analysis (Fig. 3a); however, this effect is not significant when looking at both 3B and 7A loci combined (Fig. 3b). A dominance-like effect is observed for the 7A locus on all lactones where both one or two alternative alleles increases γ-dodecalactone concentration by an average of 0.3038 mg kg^{-1}, average γ-decalactone by 1.0899 mg kg^{-1} and average γ-octalactone by 0.0057 mg kg^{-1} which are 1.8-fold, 1.49-fold and 1.34-fold increases respectively (Fig. 3a; Table S8).

In a two-way ANOVA, there was no significant interaction between the two loci for any of the lactones, and both loci additively contribute to lactone concentration (Tables S7, S8). The highest γ-decalactone and γ-octalactone concentrations are reached with two alternative alleles at the 3B locus and at least one alternative allele at the 7A locus, and these two groups produce significantly higher γ-decalactone than the rest of the allelic combinations (Table S8). Maximum γ-dodecalactone concentration is reached with at least one alternative allele at the 7A locus (Table S8).

Identification and validation of *FaECH* in γ-lactone biosynthesis

The 7A locus for lactone production discovered in GWAS was defined as 707 kbp encompassing all SNPs above the Bonferroni-corrected significance level (Fig. 4a). *Cis* eQTLs were calculated from a previous strawberry fruit expression dataset (Fan *et al.*, 2022) using the Medallion haplotype 1 genome and were filtered to retain only those on chromosome 7A (Dataset S2). Of the 76 annotated genes within the 7A locus, 11 had significant eQTL associations with AX123363242 (Dataset S3). AX123363242 had the strongest effect on the gene Fxa7A_h1_g62386, which was annotated as an Enoyl-CoA hydratase/isomerase by InterProScan (Paysan-Lafosse *et al.*, 2023) and designated as *FaECH* (Dataset S4). BLASTp analysis of *FaECH* in the Medallion haplotype 1 genome showed only one homologous gene on chromosome 7C, Fxa7C_h1_g66885 with almost no expression in the eQTL dataset (Dataset S5). *FaECH* was also the most highly upregulated over the course of fruit ripening out of the 11 genes with eQTL associations in the 7A locus (Dataset S6). Notably, individuals with the AA allele of AX123363242 displayed almost no expression of *FaECH* with increasing expression per dose of the alternative allele (Fig. 4b).

Four haplotype resolved genomes were used to investigate allelic diversity of *FaECH* in cultivated strawberry: 'Royal Royce', 'Florida Brilliance', 'FL 15.89–25' and Medallion (Fig. 4c). *FaECH* and its flanking region were found in each genome, extracted and aligned to each other, which revealed two distinct alleles for which each genome was heterozygous (Fig. 4c). All haplotypes with the A allele of AX123363242 shared the same

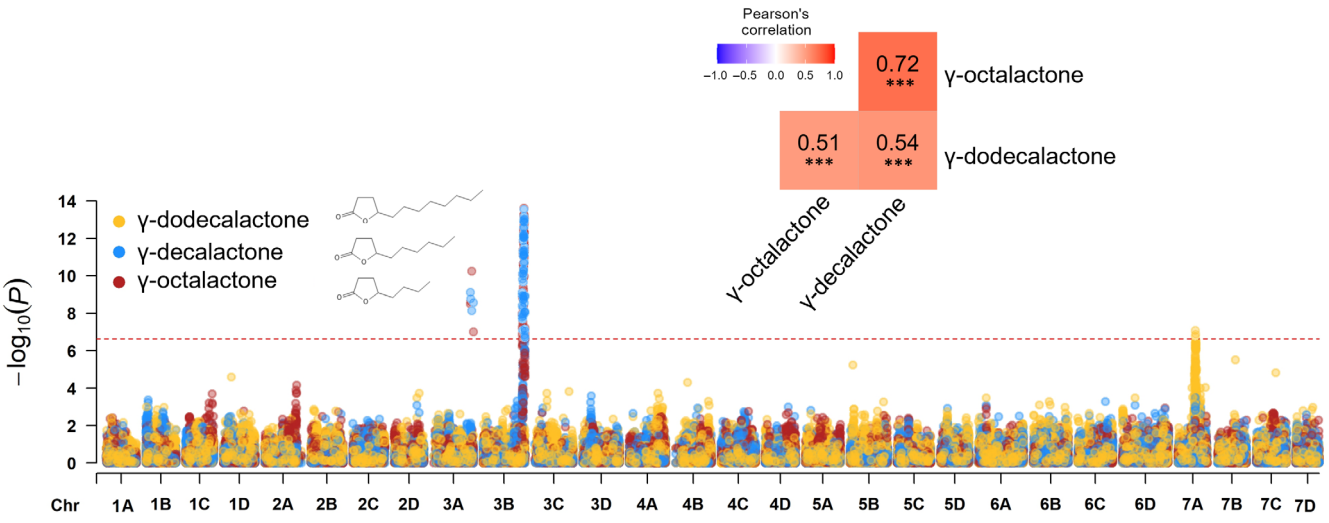

Fig. 2 Genome wide association study Manhattan plot for γ-lactones in strawberry (*Fragaria* × *ananassa*) fruit. Pearson's correlations among γ-lactones with *** indicating significance at $P < 0.0001$.

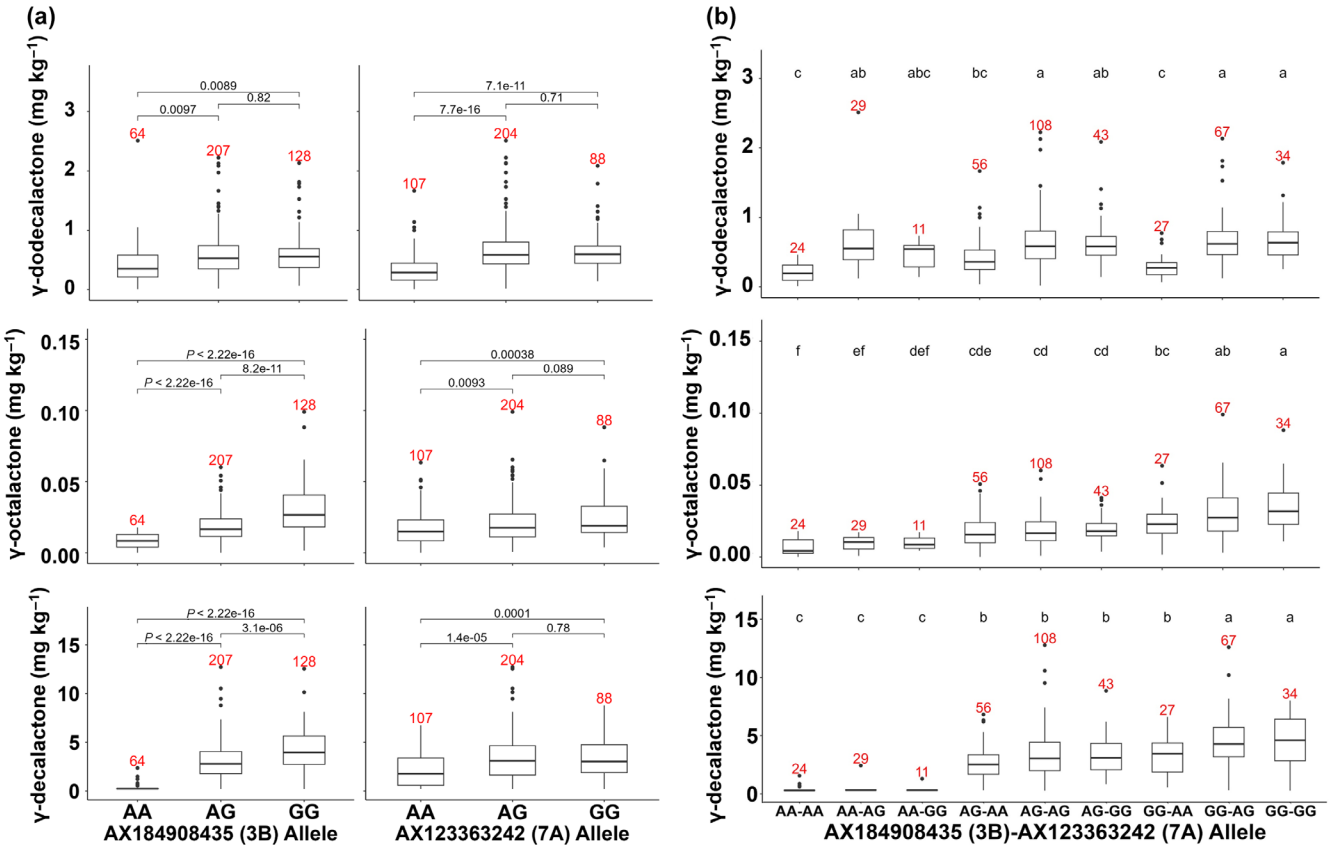

Fig. 3 SNP marker effect on γ lactones in strawberry (*Fragaria* × *ananassa*) fruit. (a) Box plots showing lactone production in each allelic class for lead SNP markers based on genome wide association study *P*-value. The left column of plots shows the effect of the chromosome 3B GWAS SNP, while the right column of boxes shows the effect of the chromosome 7A GWAS SNP. *P*-values are shown for Student's *t*-test above comparisons of allelic groups. Box plots are delimited by upper and lower quartiles with whiskers extending to the largest or smallest value within 1.5 × interquartile range from the hinges and dots represent outliers. Text indicates the number of individual data points per box. (b) Boxplots showing lactone production for 3B and 7A lead GWAS SNP markers in combination. Annotation is the same for the boxes as in panel (a), except letters indicate Tukey's HSD groups.

structure of *FaECH* and were denoted as 'reference' while the same was true for the G allele of AX123363242 and the 'alternative' allele of *FaECH* (Fig. 4c). Significant sequence divergence is observed upstream of the *FaECH* transcription start site as well as an 'A' insertion after the 157th bp resulting in a premature stop codon in the 55th amino acid in the reference allele (Fig. 4c). Also, a 568-bp deletion in the reference allele overlaps the end of the CDS (Fig. 4c). Taken together with the lack of expression in the eQTL population among individuals with two reference alleles (Fig. 4b), we conclude that sequence level differences result in a non-functional reference allele that is not expressed.

Beyond the evidence from GWAS, eQTL and ripening expression, *FaECH* has strong biochemical evidence suggested by its homology annotation as an Enoyl-CoA hydratase/isomerase enzyme, which is a key part of β-oxidation cycle that is necessary to synthesize lactones. In addition to protein homology evidence, subcellular localization prediction using DeepLoc2 (Thumuluri *et al.*, 2022) gave high confidence (0.9647) for peroxisomal localization driven by the canonical peroxisomal targeting signal-1 (PST1 site AKL (Fig. S2). BLASTp of the translated *FaECH* CDS revealed the *Arabidopsis thaliana* 3-hydroxyisobutyryl-CoA hydrolase 1 protein, CHY1, as the closest Arabidopsis homolog with 54.4% pairwise identity (Fig. S3), which has a known role in peroxisomal β-oxidation (Zolman *et al.*, 2001).

Functional validation of *FaECH* was carried out using Agrobacterium-mediated transient fruit RNAi and overexpression assays. For RNAi, Medallion was used as it produces both γ-decalactone and γ-dodecalactone. RNAi knockdown of *FaECH* showed a 3.43-fold reduction in gene expression on average, corresponding with a significant 1.76-fold average decrease in γ-dodecalactone concentration and a 1.51-fold average decrease in γ-decalactone (Fig. 5). Ember™ 'FL 20.80-4' was used for overexpression due to its lower production of lactones relative to Medallion. Overexpression of *FaECH* caused a dramatic 8323.44-fold increase in gene expression on average resulting in a significant 12.65-fold increase in γ-dodecalactone concentration as well as a 21.14-fold increase in γ-decalactone (Fig. 5).

In order to leverage the *FaECH* gene in a breeding context, a codominant PACE marker was developed from a polymorphism 666 bp away from the end of the predicted CDS at 13 022 423 bp (Medallion haplotype 1) or 12 686 848 bp on chromosome 7A ('Royal Royce') (Table S5). The PACE marker

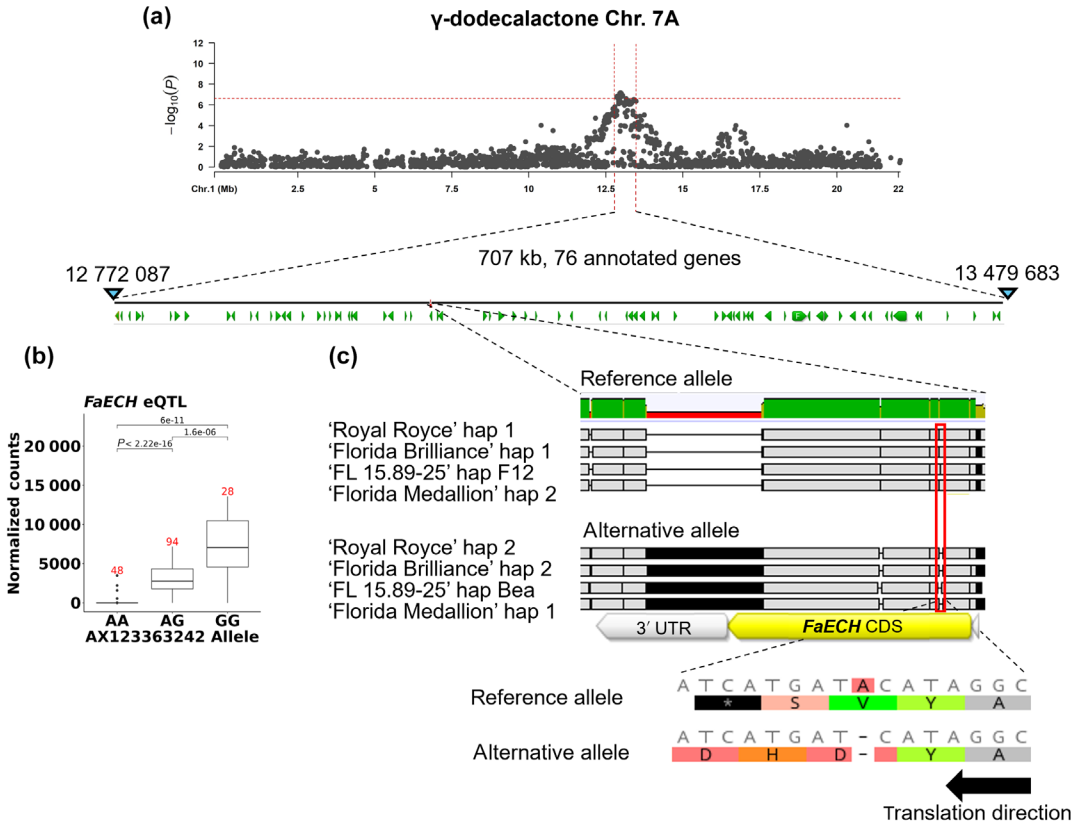

Fig. 4 Investigation of the 7A lactone locus in strawberry (*Fragaria* × *ananassa*) haplotype phased genomes. (a) Genome wide association study Manhattan plot on chromosome 7A with physical SNP marker position from the Medallion haplotype 1 genome. Expanded area shown as a continuous track with green arrows representing annotated genes. (b) Boxplot showing the expression in normalized gene counts of individuals in the eQTL dataset by AX123363242. *P*-values are shown for Student's *t*-test above comparisons of allelic groups. Box plots are delimited by upper and lower quartiles with whiskers extending to the largest or smallest value within 1.5× interquartile range from the hinges and dots represent outliers. Text indicates the number of individual data points per box. (c) Pairwise alignment of the *FaECH* locus from eight haplotypes of cultivated strawberry genomes. Light gray indicates identical nucleotide bases, black boxes indicate different nucleotide bases and black lines indicate an absence of nucleotide bases.

discriminated all three allelic classes (Fig. 6a), with high linkage to lead SNP from GWAS, AX123363242 (r^2 = 0.829). Slightly less variation in γ-dodecalactone concentration was explained by the PACE marker (r^2 = 0.087) compared to AX123363242 (r^2 = 0.097) in the individuals tested with the PACE marker (Fig. 6b).

Analysis of homologous lactone biosynthesis genes

Enzymes involved in the biosynthesis of lactones have been well-studied in peach (*P. persica*), which also belongs to the Rosaceae family. Here, we aim to use the known information about lactone biosynthesis in peach to suggest which strawberry genes could be responsible for lactone biosynthesis in strawberry. First, the suggested lactone pathway from Li *et al.* (2023) was reproduced here in Fig. 7 with some modifications. We identified nine peach genes from the literature with evidence of involvement in lactone biosynthesis (Table S9). Five genes were validated through transient gene expression and other functional studies, including two fatty acid desaturases (*PpFAD2-1*, *PpFAD2-2*) (Sánchez *et al.*, 2013; Peng *et al.*, 2023), two acyl-CoA oxidases (*PpACX1*, *PpACX5*) (Zhang *et al.*, 2017) and an alcohol acyltransferase (*PpAAT1*) (Peng *et al.*, 2020) (Table S9). Among the genes without molecular validation, there were three epoxide hydrolases (*PpEPH*, *PpEPH2* and *PpEPH3*) and a cytochrome P450 gene (Prupe.4G152700). Since no enoyl-CoA hydratase genes were suggested by the peach literature, the *FaECH* protein sequence was searched in the peach protein sequences to find its closest homolog.

BLASTp analysis of peach genes using Medallion haplotype 1 returned 112 unique genes for the nine query sequences (Dataset S7). For *PpFAD2-1*, the top BLASTp hits were homologous genes on chromosomes 6A, 6C and 6D (Fxa6A_h1_g50437, Fxa6C_h1_g56516 and Fxa6D_h1_g59235) with 87.20% identity while the next best hit was the known functional fatty acid desaturase gene, *FaFAD1*, with 71.20% identity. Fxa6A_h1_g50437, Fxa6C_h1_g56516 and Fxa6D_h1_g59235 had much lower average expression in the eQTL population with 1154, 254 and 359 normalized gene counts, respectively, compared to 16 398 normalized counts for *FaFAD1* (Dataset S7). Additionally, *FaFAD1* is highly upregulated over the course of fruit ripening (Sánchez-Sevilla *et al.*, 2014) while the other genes are downregulated (Dataset S7). Homologous genes on all subgenomes of Chromosome 5 showed pairwise identity above 89%

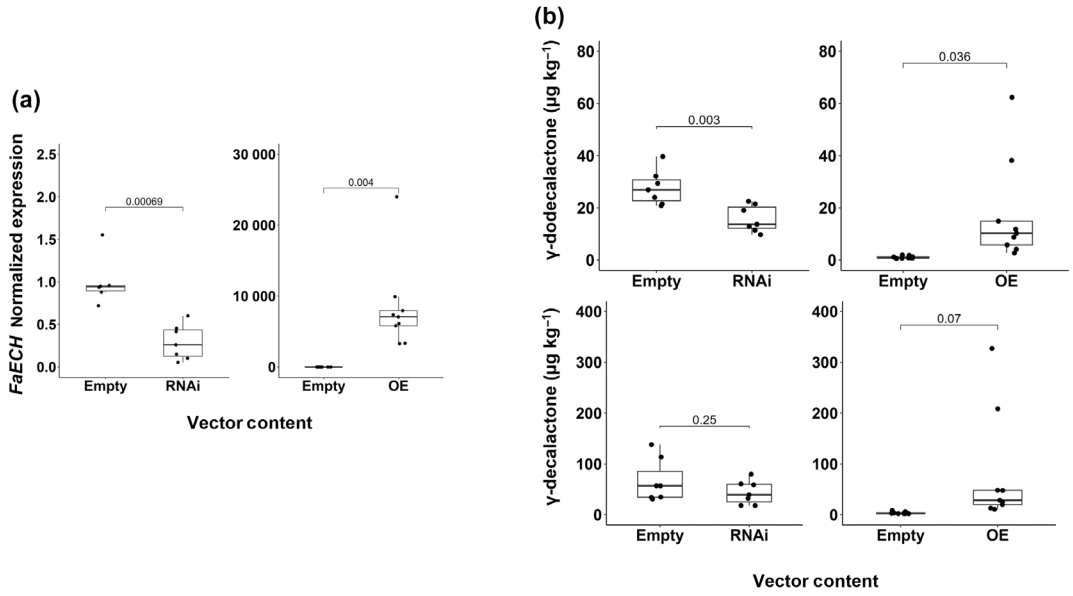

Fig. 5 RNA interference (RNAi) and overexpression strawberry (*Fragaria* × *ananassa*) fruit transient assay for *FaECH*. (a) Gene expression of *FaECH* in transiently infected strawberry fruit with average expression for the empty vector fruit set to 1. *P*-values are shown for Student's *t*-test above comparisons of allelic groups. Box plots are delimited by upper and lower quartiles with whiskers extending to the largest or smallest value within 1.5× interquartile range from the hinges and dots represent outliers. (b) Boxplots showing γ-dodecalactone and γ-decalactone concentration in fruits injected with Agrobacterium strains carrying the empty, RNAi or overexpression (OE) constructs for *FaECH*. Box plot annotations are the same as in panel (a).

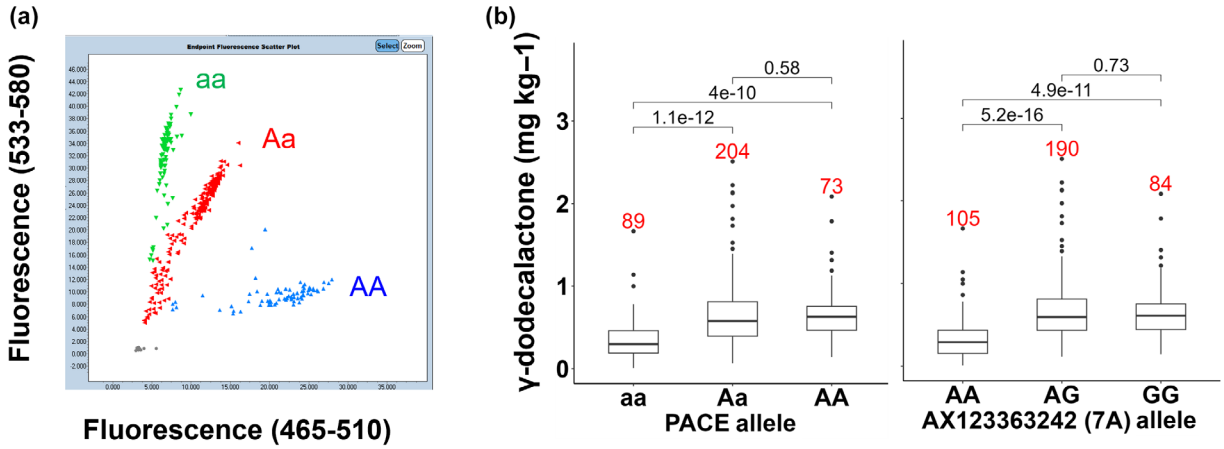

Fig. 6 PCR Allele Competitive Extension (PACE) marker analysis in strawberry (*Fragaria* × *ananassa*). (a) Endpoint genotyping clusters for the three allelic classes of the PACE marker. Gray points indicate no allele call. (b) Comparison of γ-dodecalactone production using PACE alleles or GWAS SNP alleles. The SNP allele boxplot shows only individuals from the population that were also tested with PACE markers. *P*-values are shown for Student's *t*-test above comparisons of allelic groups. Box plots are delimited by upper and lower quartiles with whiskers extending to the largest or smallest value within 1.5× interquartile range from the hinges and dots represent outliers. Text indicates the number of individual data points per box.

with *PpACX1* and 77% with *PpACX5*, moderate expression and strong ripening upregulation. The copy on chromosome 5C, Fxa5C_h1_g45829, has the highest expression level in the eQTL population (5379 normalized gene counts) and the largest increase in expression over ripening, making it a potential actor in the lactone pathway (Dataset S7). Results for *PpAAT1* showed 64% identity to Fxa5D_h1_g48006 and Fxa5A_h1_g40818 previously described as *FaAAT5* and *FaAAT6* (Saez *et al.*, 2024); however, both genes were hardly expressed in the eQTL population (Dataset S7). Fxa3C_h1_g26725 previously described as *FaAAT2* (Cumplido-Laso *et al.*, 2012), had 61% identity to *PpAAT1* and was the only markedly expressed homolog in the eQTL population (average normalized gene counts 64 101) (Dataset S7). This gene has been functionally validated to catalyze ester production (Cumplido-Laso *et al.*, 2012), but its role in lactone cyclization in strawberry fruit has not yet been

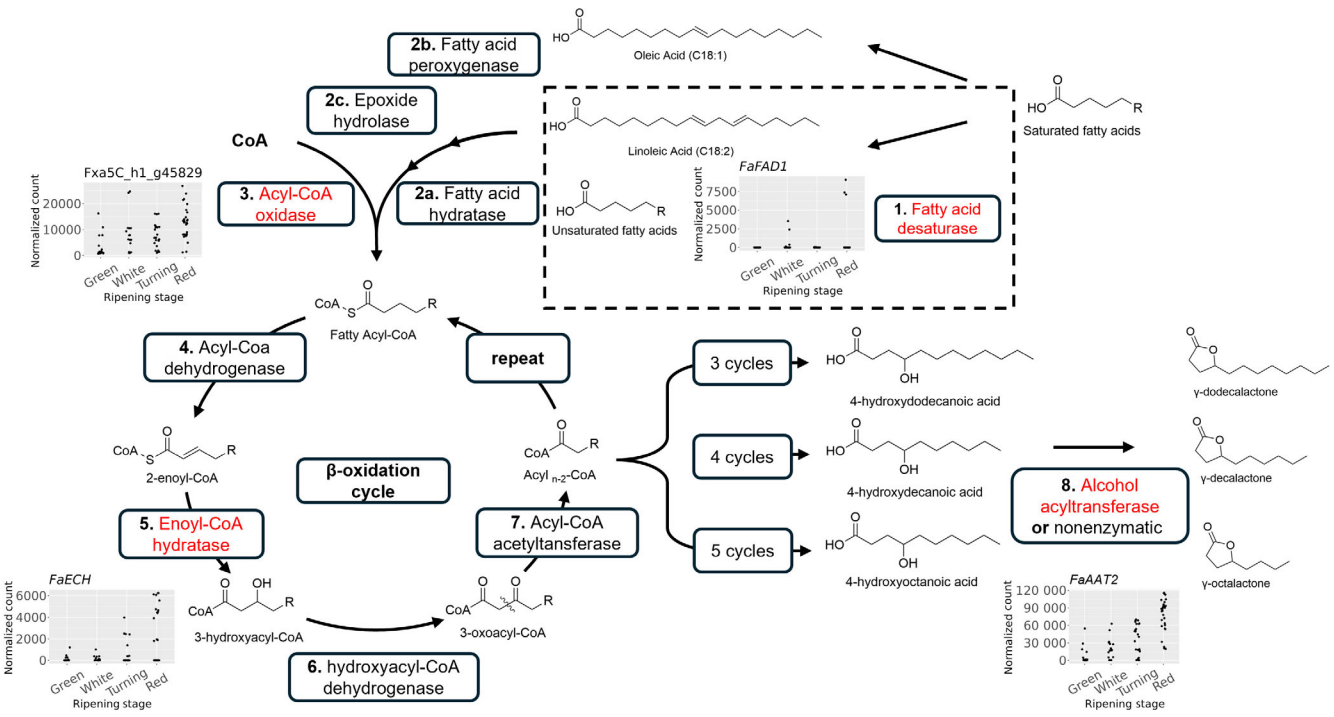

Fig. 7 Proposed biosynthesis pathway of γ-lactones in strawberry (*Fragaria* × *ananassa*). Biosynthesis enzymes highlighted in red are discussed in this study. Scatter plots display normalized gene counts across four different ripening stages for known or candidate genes in strawberry lactone biosynthesis.

Table 2 Lactone biosynthesis genes in *Fragaria* × *ananassa* with putative peach homologs.

Peach gene alias	*Prunus_persica*_v2.0.a1	Strawberry gene name	Strawberry gene alias	Average strawberry expression	Pairwise identity	Function
PpFAD2-1 or *PpFAD1B_6*	Prupe.7G076500.1	Fxa3B_h1_g24404	*FaFAD1*	16 398	71.20%	1. Fatty acid desaturase
PpACX1	Prupe.5G065100.1	Fxa5C_h1_g45829	NA	5379	89.60%	3. Acyl-CoA oxidase
NA	Prupe.2G109100.1	Fxa7A_h1_g62386	*FaECH*	2991	57.90%	5. Enoyl-CoA hydratase
PpAAT1	Prupe.5G018200.1	Fxa3C_h1_g26725	*FaAAT2*	64 101	61.30%	8. Alcohol acyltransferase

demonstrated. All three epoxide hydrolase genes, *PpEPH*, *PpEPH2* and *PpEPH3*, had different top hits in strawberry with none having high average expression in the eQTL panel and increasing expression over ripening (Dataset S7). Similarly, no hits for the CYP450 gene, Prupe.4G152700, showed high expression or ripening upregulation (Dataset S7). BLASTp search of *FaECH* in the peach genome returned multiple genes on Chromosome 2 with the highest pairwise identity at 57.9% (Table 2). Ripening upregulation of potential and known lactone biosynthesis genes are shown in Fig. 7, and more details are available in Table 2.

Lactone production in *F. vesca*

Next, to explore the diversity of γ-lactone biosynthesis in the *Fragaria* genus, we aim to analyze their production in the diploid woodland strawberry, *F. vesca*. A volatilome profile from fruits of a European *F. vesca* population was previously established (Urrutia *et al.*, 2023), identifying the same three lactones detected in the cultivated strawberry: γ-decalactone, γ-dodecalactone and γ-octalactone. In particular, lactones were quantified in two harvest years: 2016 (H16) and 2017 (H17). Using an ANOVA across both years, it was found that γ-decalactone was highly genetically controlled (ω^2_G = 0.8023), while genetic control of γ-dodecalactone was moderate (ω^2_G = 0.3983) and low for γ-octalactone (ω^2_G = 0.2569) (Urrutia *et al.*, 2023). This *F. vesca* population was also recently resequenced (Toivainen *et al.*, 2026). Thus, to elucidate the mechanisms of genetic control of lactones, we performed GWAS using both MLM and FarmCPU, given the highly structured nature of the population.

GWAS for γ-decalactone production showed a consistent signal with significant SNPs from both harvests and models centered around the end of Chromosome 3. The most significant SNP, Fvb3 1103 870, associated with the H17 harvest is located 2666 bp away from the closest *FaFAD1* homolog in *F. vesca*, FvH4_3g44270 named here *FvFAD1*, which is located at 36905750 bp (Fig. 8a; Dataset S8). Additionally, the most

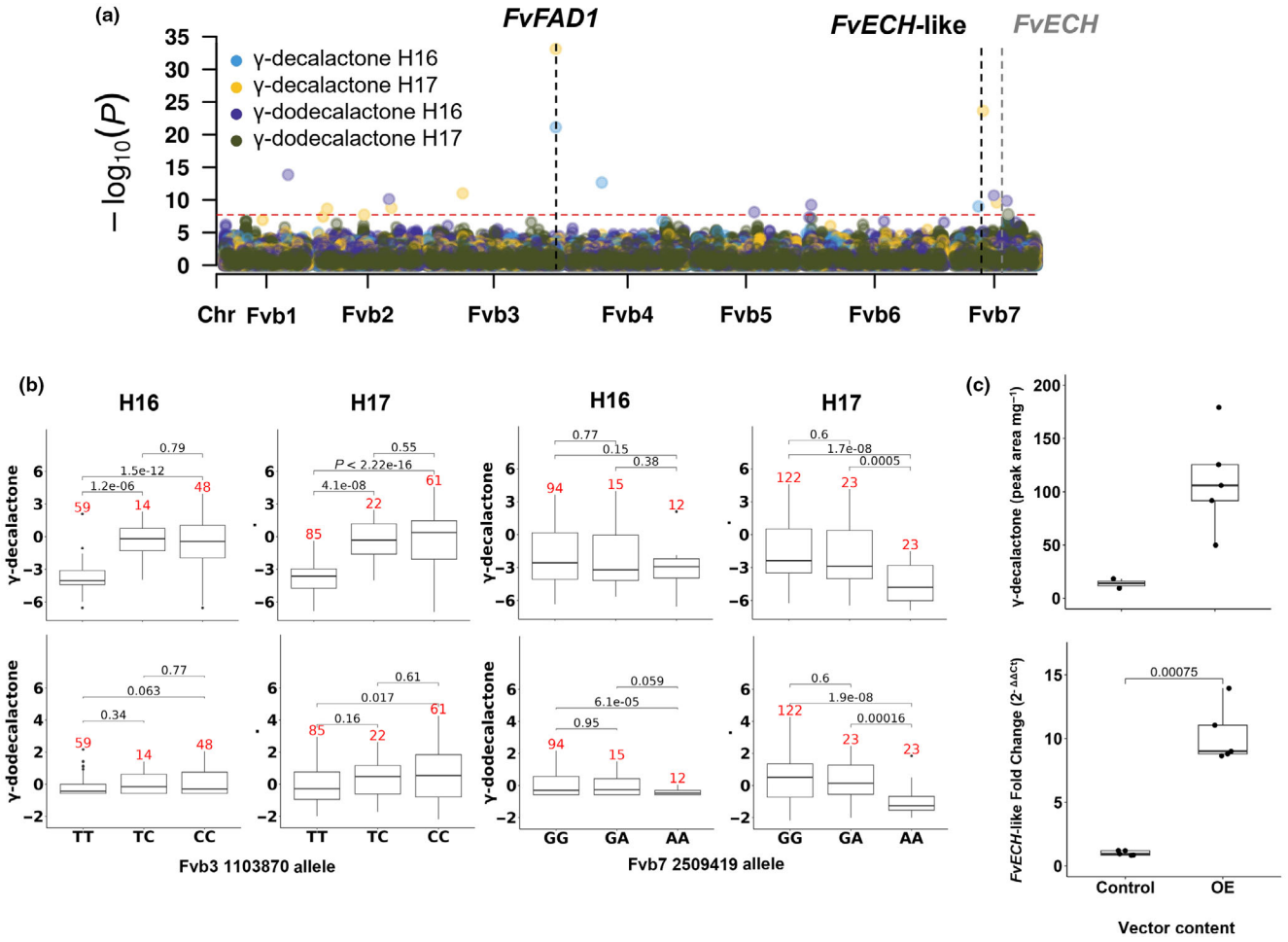

Fig. 8 Genetic analysis of lactone production in *Fragaria vesca*. (a) Manhattan plot of FarmCPU genome wide association study results with dashed black lines indicating the approximate position of discussed genes. *FvECH* is shown for reference although there was no significant SNP found nearby. (b) Box plots of γ-lactone production by lead SNP allele at Chromosome 3 and 7 loci. H16 and H17 refer to harvest seasons 2016 and 2017, respectively. *P*-values are shown for Student's *t*-test above comparisons of allelic groups. Box plots are delimited by upper and lower quartiles with whiskers extending to the largest or smallest value within 1.5× interquartile range from the hinges and dots represent outliers. Text indicates the number of individual data points per box. (c) Volatile quantification and gene expression for *FvECH*-like transient overexpression fruit assay. Boxplot annotations are the same as in panel (b).

significant signal across the rest of the chromosomes was identified in H17 by the FarmCPU model toward the middle of chromosome 7, Fvb7 2509 419. This variant is located 15 294-bp upstream of a predicted Enoyl-CoA hydratase/isomerase, HIBYL-CoA-H type gene, FvH4_7g09160 (hereafter referred to as *FvECH*-like) at 8 903 798 bp (Fig. 8a; Dataset S8). Notably, *FvECH*-like is distinct from the *FaECH* ortholog in *F. vesca* (FvH4_7g16010; *FvECH*). The approximate position of FvH4_7g16010 or *FvECH* is shown in Fig. 8 to compare with *FvECH*-like. The physical location of *FvECH* is 13 774 438, which is 935 644 bp from the closest significant SNP for any of the γ-lactones. GWAS for γ-dodecalactone production did not yield any significant signals using the standard MLM model. The FarmCPU model showed multiple signals across the genome with a few on Chromosome 7; however, none overlapping across both harvests or with signals from the γ-decalactone GWAS. No significant associations were detected for γ-octalactone on any chromosome.

The effect of the lead SNP on chromosome 3, Fvb3 1103 870, on γ-decalactone production is similar to the effect of the *FaFAD1* locus in *F.* × *ananassa* with *F. vesca* genotypes that have the TC or CC alleles producing significantly more γ-decalactone than the group with the TT allele in both harvests (Fig. 8b). The lead SNP on Chromosome 7, Fvb7 2509 419, affects γ-decalactone significantly, with groups that have the GG and GA alleles producing significantly more than the group with the AA allele in H17 but no significant difference in H16. Although SNP Fvb7 2509 419 was not significant in GWAS for γ-dodecalactone production, individuals with GG and GA alleles produce significantly more than those with the AA allele in H17 (Fig. 8b).

Investigation of the *FvECH*-like CDS in the *F. vesca* population revealed two polymorphisms, which segregate near perfectly with SNP Fvb7 2509419. Among 26 individuals with the AA allele at Fvb7 2509419, all encode a truncated *FvECH*-like gene

resulting from a C to T transition in the 109th bp (Fig. S4). The 144 individuals with the GG allele at Fvb7 2509419 almost exclusively encode the full-length *FvECH*-like CDS with the exception of three genotypes (Fig. S5). To determine whether *FvECH*-like affects γ-lactone production, functional validation of the full-length CDS allele was carried out using a transient fruit overexpression assay in *F. vesca* resulting in a 10.28-fold increase in *FvECH*-like expression for the overexpression samples (Fig. 8c). In three of the five control samples, γ-decalactone was undetected with no peak of the target ion (Figs 8, S6) while γ-decalactone was detected in all five overexpression samples with 8.43-fold higher abundance than the two control samples with detectable lactones, supporting the role of *FvECH*-like in the biosynthesis of this lactone.

Comparative analysis of lactone genes in *F. vesca* and *F. × ananassa*

Despite pairwise identity of 93.1% between *FaFAD1* and *FvFAD1* protein sequences, they exist on nonhomologous chromosomes 3A and 3B, respectively (subgenome A, not B, is most closely related to *F. vesca*) and exhibit significant genomic sequence differences. Alignment of the *FvFAD1* region on *F. vesca* Chromosome 3 to the *FaFAD1* region on chromosome 3B revealed a *c.* 661-bp intron present in *FaFAD1* but not *FvFAD1* (Fig. S7a). Outside of the predicted CDS regions and about 200-bp upstream of the gene annotation, sequence alignment was extremely poor, highlighting the different chromosomal context of the two genes. Alignment of the *FvFAD1* region to the homologous region on *F. × ananassa* chromosome 3A shows no annotated FAD gene in the *F. × ananassa*, but a 3500-bp deletion encompassing most of where the *FvFAD1* gene body would be (Fig. S7b). These findings suggest that the function of the *FAD1* gene from the *F. vesca* ancestor has been lost in *F. × ananassa*, and *FaFAD1* was instead acquired from a different wild ancestor during polyploidization and carries out a similar function as it relates to lactone production. In cultivated strawberry, it is well-established that a large deletion in the *FaFAD1* region completely disrupts gene expression resulting in little γ-decalactone production (Oh *et al.*, 2021). In the *F. vesca* population, we aimed to identify sequence variation in the *FvFAD1* CDS associated with the SNP Fvb3 1103 870 identified through GWAS. Among individuals with the CC allele at Fvb3 1103870, all 66 genotypes have alleles of *FvFAD1* that are highly similar however, a G to A substitution at the 846th nucleotide that leads to an early stop codon is present in 44/66 individuals (Fig. S8). Despite this difference, there is no difference or a slight increase in γ-lactone production for individuals that have the predicted early stop codon (Fig. S9a). Among individuals with the TT allele, 64/101 encode a truncated allele with a different early stop codon resulting from a C to G substitution at the 153rd position (Fig. S10). The rest, 37/101 have the nontruncated allele. There is no difference or a slight increase in γ-lactone production for individuals that do not have this predicted early stop codon (Fig. S9b). We are not able to conclusively identify major mutations in the *FvFAD1* CDS that are linked to the GWAS SNP allele.

We also compared *FaECH* and its *F. vesca* homolog, FvH4_7g16010 (*FvECH*). They have high similarity in the 3500-bp region encompassing the upstream, gene body and downstream genomic sequence (Fig. S7d). These differences result in highly similar protein translations between the two species; however, most differences lie in the 672-bp region upstream of the annotated gene. In a previous report comparing three *F. vesca* genotypes at three ripening stages, *FvECH* (gene43078, as previously annotated in the v.2.0.a2 *F. vesca* genome version) was not expressed (Härtl *et al.*, 2017), whereas both *FvECH-like* and *FvFAD1* (gene05329 and gene24414) increased in their expression during ripening (Dataset S9). Given that *FvECH* was not expressed over fruit ripening, unlike the other lactone biosynthesis genes, and no significant GWAS SNPs were found near *FvECH* in this study, it likely does not play a major role in lactone biosynthesis in *F. vesca*.

FvECH-like and its *F. × ananassa* homolog, Fxa7A_h1_g61875 (*FaECH-like*), show high conservation of gene CDS and the surrounding 2000 bp (Fig. S7c). Additionally, *FaECH-like* showed 1.47-fold higher average expression in the eQTL population and increased ripening upregulation compared to *FaECH* (Dataset S5). We also detected a strong *cis*-eQTL for *FaECH-like* that shows a similar dosage-like effect for *FaECH*-like expression as the *cis*-eQTL for *FaECH* discussed earlier (Fig. S11). Finally, there are no significant associations between any of the *F. × ananassa* lactone phenotypes and SNPs near *FaECH-like*, which is located at 8412725 bp on chromosome 7A. This evidence suggests that contrary to *FvECH*-like in *F. vesca*, *FaECH-like* may not play a major role in lactone biosynthesis in *F. × ananassa*.

Discussion

In this study, we quantified γ-lactones in a large, cultivated strawberry breeding population, facilitating GWAS and the discovery of a novel locus modulating concentrations of γ-decalactone, γ-dodecalactone and γ-octalactone. Genome assembly of a high lactone cultivar, Medallion™ 'FL 16.30-128', allowed for remapping of eQTL data, resulting in the detection of a strong *cis*-eQTL for a putative enoyl-CoA hydratase/isomerase gene (*FaECH*). Through transient fruit assays, we showed the ability of *FaECH* to affect γ-dodecalactone and γ-decalactone accumulation in strawberry fruit. A PACE marker was developed to select for the functional allele of *FaECH* for the breeding of more flavorful strawberries. We also used the Medallion genome annotations as a reference for BLASTp analysis of peach genes involved in lactone biosynthesis, which clarified the roles of *FaECH* and *FaFAD1* in the strawberry lactone pathway. Using a diverse population of wild, diploid strawberry, *F. vesca*, we performed GWAS and identified two loci associated with γ-decalactone production. We determined the effect of the Chromosome 3 locus was likely due to variation in the *FvFAD1* gene, exhibiting similar genetic control of lactone production to the *FaFAD1* gene in *F. × ananassa*. The Chromosome 7 locus is linked to an enoyl-CoA hydratase/isomerase gene, *FvECH*-like, which is nonhomologous with *FaECH*. Transient overexpression of *FvECH*-like in *F. vesca* fruit increased γ-decalactone concentration.

Observing the genetic control of lactones in *F.* × *ananassa* and integrating knowledge of lactone biosynthesis from peach, we outlined the known lactone biosynthesis pathway in strawberry and suggested other potential enzymes based on homology and expression patterns (Fig. 7; Table 2). While *FaFAD1* is the main enzyme that generates desaturated fatty acids that are further modified to create lactones, it does not seem to be responsible for increasing γ-dodecalactone content, given the lack of GWAS signal on chromosome 3B and subsequent SNP effects (Figs 1, 3b). An explanation for this is that the main precursor for γ-dodecalactone is oleic acid as has been demonstrated in yeast (Haffner & Tressl, 1996), and given that *FaFAD1* encodes an omega-6 fatty acid desaturase (Chambers *et al.*, 2014; Sánchez-Sevilla *et al.*, 2014), it converts oleic acid into linoleic acid. In peach, the homologous fatty acid desaturase, *PpFAD2-1*, was demonstrated to desaturate oleic acid to linoleic acid, resulting in the dramatic increase in the production of the shorter lactones γ-decalactone and γ-octalactone (Sánchez *et al.*, 2013; Peng *et al.*, 2023). Given the likely role of *FaECH* in β-oxidation, its effect on the accumulation of all the γ-lactones is consistent (Figs 3a, 5) since all γ-lactone precursors undergo β-oxidation regardless of carbon length. Further validation of *FaECH* and identification of its substrates will pave the way for identifying bottlenecks that may limit lactone biosynthesis in *F.* × *ananassa*. In summary, *FaFAD1* and *FaECH* are two major genes responsible for γ-lactone production with natural allelic variation that can be selected on to improve strawberry flavor.

Leveraging a sizeable European diversity panel of *F. vesca* (Urrutia *et al.*, 2017; Jiménez-Muñoz *et al.*, 2025; Toivainen *et al.*, 2026), we were able to conduct GWAS and compare the genetic control of γ-lactones between the diploid and octoploid strawberry species. Remarkably, the strongest genetic effect came from Chromosome 3 in *F. vesca* (Fig. 8), just 4812 bp away from the closest homolog to *FaFAD1* referred to here as *FvFAD1*. While *FvFAD1* has not been validated in transient fruit assays, it appears likely that it carries out the same role in *F. vesca* that *FaFAD1* does in *F.* × *ananassa*. We also observed the corresponding region for *FvFAD1* on the vesca-like chromosome 3A in the octoploid genome contained a large deletion and no annotated *FAD* gene. Perhaps, this is as a result of evolutionary forces retaining only one functional *FAD* gene at the octoploid level, or that the final polyploidization event between an ancestral hexaploid and ancestral *F. vesca* (Fan *et al.*, 2025) occurred with an ancestral *F. vesca* lacking functional *FvFAD1* alleles. *F. vesca* could potentially be used to introgress an additional functional *FAD1* allele into *F.* × *ananassa* to increase γ-decalactone production similar to the use of the diploid *F. nilgerrensis* to produce the high-lactone decaploid cultivar 'Tokun' (Noguchi, 2011). Beyond *FvFAD1*, we found that *FvECH*-like modulates γ-decalactone production in *F. vesca* while the *FaECH* homolog *FvECH* does not. *FvECH* was not expressed in a previous ripening upregulation dataset in *F. vesca* (Dataset S9) (Härtl *et al.*, 2017), which may explain why it does not show the same effect as *FaECH* does in *F.* × *ananassa* despite the high similarity between the two genes (Fig. S7). Expression of the *FvECH*-like homolog in *F.* × *ananassa*, *FaECH-like* was high and affected by a *cis*-eQTL (Fig. S11). Despite this, there were no significant GWAS SNPs located nearby suggesting that *FaECH-like* does not affect lactone accumulation in *F.* × *ananassa*.

Beyond the proposed function of *FaECH* and *FvECH*-like in lactone biosynthesis, their activity in β-oxidation and upregulation in late fruit ripening suggests involvement in core metabolic processes. While the role of β-oxidation in fruit ripening is not well- characterized, an increase in β-oxidation activity, including *ECH* gene expression and protein abundance, has been demonstrated concurrent with ripening in pepper and grape (González-Gordo *et al.*, 2022; Olmedo *et al.*, 2023). Increasing β-oxidation activity over ripening could be important for efficient degradation of fatty acids produced from membrane turnover (Marano *et al.*, 1993; Graham & Eastmond, 2002). The resulting acetyl-CoA produced by β-oxidation of fatty acids may be funneled into energy production (Graham & Eastmond, 2002) or be used by *FaAAT* enzymes for creation of esters characteristic in ripe strawberry fruit (Cumplido-Laso *et al.*, 2012; Saez *et al.*, 2024). Additionally, β-oxidation is necessary for the production of jasmonic acid, indole acetic acid and salicylic acid (Nyathi & Baker, 2006), of which only salicylic acid increases over the course of strawberry ripening (Kim *et al.*, 2019). The role of *FaECH* and *FvECH*-like in ripening physiology warrants further investigation.

With this research, we contribute greater understanding of γ-lactone production in *F.* × *ananassa* with a practical PACE marker for *FaECH* to be deployed in breeding programs. We suggest that selecting for at least one functional allele of *FaECH* and two of *FaFAD1* is optimal to generate breeding lines with perceptible levels that impact flavor. Medallion™ 'FL 16.30-128' is a released cultivar with excellent flavor that has the optimal allelic combination resulting from the use of marker-assisted selection for *FaFAD1* (Dalid *et al.*, 2025). We also contribute extensive discoveries of the genetic control over γ-lactones in *F. vesca* and define two major genes in *FvFAD1* and *FvECH*-like. This knowledge can be used to develop molecular breeding tools to improve *F. vesca* flavor or leverage the wider germplasm pool to introgress novel alleles into cultivated varieties. While flavor is a quantitative trait that depends on sugars, acids and many volatile compounds, increasing concentrations of specific volatile compounds is necessary to create unique flavor profiles that consumers will enjoy.

Acknowledgements

We thank the personnel of the University of Florida strawberry breeding program for maintenance of *F.* × *ananassa* plants and collection of fruit samples. Work on *F.* × *ananassa* was supported by grants awarded to Vance M. Whitaker and Seonghee Lee from the USDA National Institute of Food and Agriculture (NIFA) Specialty Crops Research Initiative Awards #2017-51181-26833 and #2022–51181-38328; and the Florida Strawberry Research and Education Foundation (FSREF). Work on *F. vesca* was supported by grants awarded by David Posé (European Research Council, grant no. ERC-2014-StG 638134, and Ministerio de Ciencia e Innovación, grant nos. PID2021-123677OB-I00 and PID2024-155863OB-I00) and Carmen Martín-Pizarro (Junta

de Andalucía grant nos. UMA20-FEDERJA-093 and POST-DOC21_00893).

Competing interests

None declared.

Author contributions

MP, SL, VW and ZF designed the study. MP performed bioinformatic and statistical analyses on *F.* × *ananassa* data. XS, J-PA and YW performed chemical method development and analyses for *F.* × *ananassa*. DP, TH, CM-P and MGB provided *F. vesca* data and analysis. MP and MGB performed fruit transient assays. MP, SL and VW composed the manuscript. All authors contributed edits and reviewed the final manuscript.

ORCID

Maria G. Blanco https://orcid.org/0009-0000-0755-1065
Timo Hytönen https://orcid.org/0000-0002-5231-4031
Seonghee Lee https://orcid.org/0000-0002-5190-8014
Carmen Martín-Pizarro https://orcid.org/0000-0001-9305-3096
Mark Porter https://orcid.org/0009-0008-0834-0161
David Posé https://orcid.org/0000-0003-3332-4661
Xuebo Song https://orcid.org/0000-0001-9491-3203
Yu Wang https://orcid.org/0000-0002-2003-270X
Vance M. Whitaker https://orcid.org/0000-0002-2172-3019

Data availability

The phased genome assembly for Florida Medallion is available via GenBank under GCA_049309125.1 for Haplotype 1 and GCA_049308895.1 for Haplotype 2. *Fragaria vesca* metabolomic data are available in Urrutia *et al.* (2023) at doi: 10.1111/tpj.16404. *Fragaria vesca* variant data are available from Toivainen *et al.* (2026) at doi: 10.1038/s42003-026-09539-5 through the Dryad repository Dryad. doi: 10.5061/dryad.8cz8w9h43.

References

Ahmed F, Senthil-Kumar M, Dai X, Ramu VS, Lee S, Mysore KS, Zhao PX. 2020. pssRNAit: a web server for designing effective and specific plant siRNAs with genome-wide off-target assessment. *Plant Physiology* **184**: 65–81.

Alonge M, Lebeigle L, Kirsche M, Jenike K, Ou S, Aganezov S, Wang X, Lippman ZB, Schatz MC, Soyk S. 2022. Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biology* **23**: 258.

Browning BL, Zhou Y, Browning SR. 2018. A one-penny imputed genome from Next-Generation reference panels. *American Journal of Human Genetics* **103**: 338–348.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Chambers AH, Pillet J, Plotto A, Bai J, Whitaker VM, Folta KM. 2014. Identification of a strawberry flavor gene candidate using an integrated genetic-genomic-analytical chemistry approach. *BMC Genomics* **15**: 217.

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. 2021. Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. *Nature Methods* **18**: 170–175.

Cumplido-Laso G, Medina-Puche L, Moyano E, Hoffmann T, Sinz Q, Ring L, Studart-Wittkowski C, Caballero JL, Schwab W, Muñoz-Blanco J *et al.* 2012. The fruit ripening-related gene FaAAT2 encodes an acyl transferase involved in strawberry aroma biogenesis. *Journal of Experimental Botany* **63**: 4275–4290.

Dalid C, Osorio LF, Peres NA, Lee S, Whitaker VM, Plotto A, Bai J. 2025. Florida Medallion™ 'FL 16.30-128' Strawberry. *HortScience* **60**: 1032–1036.

Damgaard MV, Treebak JT. 2022. Protocol for qPCR analysis that corrects for cDNA amplification efficiency. *STAR Protocols* **3**: 101515.

Delaneau O, Ongen H, Brown AA, Fort A, Panousis NI, Dermitzakis ET. 2017. A complete tool set for molecular QTL discovery and analysis. *Nature Communications* **8**: 15452.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

Fan Z, Hasing T, Johnson TS, Garner DM, Schwieterman ML, Barbey CR, Colquhoun TA, Sims CA, Resende MFR, Whitaker VM. 2021. Strawberry sweetness and consumer preference are enhanced by specific volatile compounds. *Horticulture Research* **8**: 66.

Fan Z, Liston A, Soltis DE, Soltis PS, Ashman T-L, Hummer KE, Whitaker VM. 2025. Homoploid hybridization adds clarity to the origins of octoploid strawberries. *Proceedings of the National Academy of Sciences, USA* **122**: e2502814122.

Fan Z, Plotto A, Bai J, Whitaker VM. 2021. Volatiles influencing sensory attributes and Bayesian modeling of the soluble solids-sweetness relationship in strawberry. *Frontiers in Plant Science* **12**: 640704.

Fan Z, Tieman DM, Knapp SJ, Zerbe P, Famula R, Barbey CR, Folta KM, Amadeu RR, Lee M, Oh Y *et al.* 2022. A multi-omics framework reveals strawberry flavor genes and their regulatory elements. *New Phytologist* **236**: 1089–1107.

Goel M, Sun H, Jiao W, Schneeberger K. 2019. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biology* **20**: 277.

González-Gordo S, Palma JM, Corpas FJ. 2022. Peroxisomal proteome mining of sweet pepper (*Capsicum annuum* L.) fruit ripening through whole isobaric tags for relative and absolute quantitation analysis. *Frontiers in Plant Science* **13**: 893376.

Graham IA, Eastmond PJ. 2002. Pathways of straight and branched chain fatty acid catabolism in higher plants. *Progress in Lipid Research* **41**: 156–181.

Haffner T, Tressl R. 1996. Biosynthesis of (*R*)-γ-Decanolactone in the Yeast *Sporobolomyces odorus*. *Journal of Agricultural and Food Chemistry* **44**: 1218–1223.

Hardigan MA, Feldmann MJ, Lorant A, Bird KA, Famula R, Acharya C, Cole G, Edger PP, Knapp SJ. 2019. Genome synteny has been conserved among the octoploid progenitors of cultivated strawberry over millions of years of evolution. *Frontiers in Plant Science* **10**: 1789.

Hardigan MA, Feldmann MJ, Pincot DDA, Famula RA, Vachev MV, Madera MA, Zerbe P, Mars K, Peluso P, Rank D *et al.* 2021. Blueprint for phasing and assembling the genomes of heterozygous polyploids: application to the octoploid genome of strawberry. *bioRxiv*. doi: 10.1101/2021.11.03.467115.

Härtl K, Denton A, Franz-Oberdorf K, Hoffmann T, Spornraft M, Usadel B, Schwab W. 2017. Early metabolic and transcriptional variations in fruit of natural white-fruited *Fragaria vesca* genotypes. *Scientific Reports* **7**: 45113.

Jang YJ, Han H, Lee S. 2024. Agrobacterium-mediated transient expression methods to validate gene functions in strawberry (*F.* × *ananassa*). *Plants* **13**: 3290.

Jiménez-Muñoz R, Urrutia M, Meco V, Rambla JL, Toivainen T, Pérez-Hedo M, Sánchez-Sevilla JF, Urbaneja A, Salas JJ, Hytönen T *et al.* 2025. Uncovering the genetic basis of fruit volatiles in *Fragaria vesca* through GWAS reveals FvJMT2 as a methyl benzoate biosynthesis gene with insect-repellent function. *bioRxiv*. doi: 10.1101/2025.11.18.689018.

Kim J, Lee JG, Hong Y, Lee EJ. 2019. Analysis of eight phytohormone concentrations, expression levels of ABA biosynthesis genes, and ripening-

related transcription factors during fruit development in strawberry. *Journal of Plant Physiology* **239**: 52–60.

Li X, Gao P, Zhang C, Xiao X, Chen C, Song F. 2023. Aroma of peach fruit: a review on aroma volatile compounds and underlying regulatory mechanisms. *International Journal of Food Science and Technology* **58**: 4965–4979.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

Marano MR, Serra EC, Orellano EG, Carrillo N. 1993. The path of chromoplast development in fruits and flowers. *Plant Science* **94**: 1–17.

Noguchi Y. 2011. "Tokun": a new decaploid interspecific hybrid strawberry having the aroma of the wild strawberry. *Journal of Japan Association on Odor Environment* **42**: 122–128.

Nyathi Y, Baker A. 2006. Plant peroxisomes as a source of signalling molecules. *Biochimica et Biophysica Acta, Molecular Cell Research* **1763**: 1478–1495.

Oh Y, Barbey CR, Chandra S, Bai J, Fan Z, Plotto A, Pillet J, Folta KM, Whitaker VM, Lee S. 2021. Genomic characterization of the fruity aroma gene, FaFAD1, reveals a gene dosage effect on γ-decalactone production in strawberry (*Fragaria* × *ananassa*). *Frontiers in Plant Science* **12**: 639345.

Olmedo P, Vidal J, Ponce E, Defilippi BG, Pérez-Donoso AG, Meneses C, Carpentier S, Pedreschi R, Campos-Vargas R. 2023. Proteomic and low-polar metabolite profiling reveal unique dynamics in fatty acid metabolism during flower and berry development of table grapes. *International Journal of Molecular Sciences* **24**: 15360.

Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, Bileschi ML, Bork P, Bridge A, Colwell L *et al.* 2023. InterPro in 2022. *Nucleic Acids Research* **51**: D418–D427.

Peng B, Gu Z, Zhou Y, Ning Y, Xu H, Li G, Ni Y, Sun P, Xie Z, Shi S *et al.* 2023. Potential role of fatty acid desaturase 2 in regulating peach aroma formation. *Postharvest Biology and Technology* **204**: 112473.

Peng B, Yu M, Zhang B, Xu J, Ma R. 2020. Differences in PpAAT1 activity in high- and low-aroma peach varieties affect γ-decalactone production. *Plant Physiology* **182**: 2065–2080.

Pi M, Gao Q, Kang C. 2019. Transient expression assay in strawberry fruits. *Bio-Protocol* **9**: e3249.

Porter M, Fan Z, Lee S, Whitaker VM. 2023. Strawberry breeding for improved flavor. *Crop Science* **63**: 1949–1963.

Pott DM, Vallarino JG, Osorio S. 2021. Profiling volatile compounds in blackcurrant fruit using headspace solid-phase microextraction coupled to gas chromatography-mass spectrometry. *Journal of Visualized Experiments* **172**: e62421. doi: 10.3791/62421.

Rey-Serra P, Mnejja M, Monfort A. 2022. Inheritance of esters and other volatile compounds responsible for the fruity aroma in strawberry. *Frontiers in Plant Science* **13**: 959155.

Saez D, Rodríguez-Arriaza F, Urra G, Fabi JP, Hormazábal-Abarza F, Méndez-Yáñez A, Castro E, Bustos D, Ramos P, Morales-Quintana L. 2024. Unraveling the key step in the aroma puzzle: insights into alcohol acyltransferases in strawberries. *Plant Physiology and Biochemistry* **212**: 108668.

Sánchez G, Venegas-Calerón M, Salas JJ, Monforte A, Badenes ML, Granell A. 2013. An integrative "omics" approach identifies new candidate genes to impact aroma volatiles in peach fruit. *BMC Genomics* **14**: 343.

Sánchez-Sevilla JF, Cruz-Rus E, Valpuesta V, Botella MA, Amaya I. 2014. Deciphering gamma-decalactone biosynthesis in strawberry fruit using a combination of genetic mapping, RNA-Seq and eQTL analyses. *BMC Genomics* **15**: 218.

Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S *et al.* 2022. Database resources of the national center for biotechnology information. *Nucleic Acids Research* **50**: D20–D26.

Schöttler M, Boland W. 1996. Biosynthesis of dodecano-4-lactone in ripening fruits: crucial role of an epoxide-hydrolase in enantioselective generation of aroma components of the nectarine (*Prunus persica* var. *nucipersica*) and the strawberry (*Fragaria ananassa*). *Helvetica Chimica Acta* **79**: 1488–1496.

Schwieterman ML, Colquhoun TA, Jaworski EA, Bartoshuk LM, Gilbert JL, Tieman DM, Odabasi AZ, Moskowitz HR, Folta KM, Klee HJ *et al.* 2014. Strawberry flavor: diverse chemical compositions, a seasonal influence, and effects on sensory perception. *PLoS ONE* **9**: e88446.

Song X, Porter ME, Whitaker VM, Lee S, Wang Y. 2023. Identification of ethyl vanillin in strawberry (*Fragaria* × *ananassa*) using a targeted metabolomics strategy: From artificial to natural. *Food Chemistry: X* **20**: 100944.

Thumuluri V, Almagro Armenteros JJ, Johansen AR, Nielsen H, Winther O. 2022. DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Research* **50**: W228–W234.

Toivainen T, Salonen JS, Kirshner J, Lembinen S, De Kort H, Lyyski A, Edger PP, Hilmarsson HS, Hallsson JH, Sargent DJ *et al.* 2026. The Late Quaternary climate impact on the genome of the woodland strawberry (*Fragaria vesca*), a perennial herb. *Communications Biology* **9**: 263. doi: 10.1038/s42003-026-09539-5.

Ulrich D, Kecke S, Olbricht K. 2018. What do we know about the chemistry of strawberry aroma? *Journal of Agricultural and Food Chemistry* **66**: 3291–3301.

Ulrich D, Olbricht K. 2016. A search for the ideal flavor of strawberry – comparison of consumer acceptance and metabolite patterns in *Fragaria* × *ananassa* Duch. *Journal of Applied Botany and Food Quality* **89**: 223–234. doi: 10.5073/jabfq.2016.089.029.

Urrutia M, Meco V, Rambla JL, Martín-Pizarro C, Pillet J, Andrés J, Sánchez-Sevilla JF, Granell A, Hytönen T, Posé D. 2023. Diversity of the volatilome and the fruit size and shape in European woodland strawberry (*Fragaria vesca*). *The Plant Journal* **116**: 1201–1217.

Urrutia M, Rambla JL, Alexiou KG, Granell A, Monfort A. 2017. Genetic analysis of the wild strawberry (*Fragaria vesca*) volatile composition. *Plant Physiology and Biochemistry* **121**: 99–117.

Wang A-H, Ma H-Y, Zhang B-H, Mo C-Y, Li E-H, Li F. 2022. Transcriptomic and metabolomic analyses provide insights into the formation of the peach-like aroma of *Fragaria nilgerrensis* Schlecht. fruits. *Genes* **13**: 1285.

Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, Yuan X, Zhu M, Zhao S, Li X *et al.* 2021. rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated Tool for Genome-wide Association Study. *Genomics, Proteomics & Bioinformatics / Beijing Genomics Institute* **19**: 619–628.

Zhang L, Li H, Gao L, Qi Y, Fu W, Li X, Zhou X, Gao Q, Gao Z, Jia H. 2017. Acyl-CoA oxidase 1 is involved in γ-decalactone release from peach (*Prunus persica*) fruit. *Plant Cell Reports* **36**: 829–842.

Zolman BK, Monroe-Augustus M, Thompson B, Hawes JW, Krukenberg KA, Matsuda SP, Bartel B. 2001. chy1, an Arabidopsis mutant with impaired beta-oxidation, is defective in a peroxisomal beta-hydroxyisobutyryl-CoA hydrolase. *The Journal of Biological Chemistry* **276**: 31037–31046.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Dataset S1 Medallion haplotype 1 genome physical positions for Axiom™ Strawberry FanaSNP 50 k SNP markers.

Dataset S2 *Fragaria* × *ananassa cis* eQTL results for all genes on chromosome 7A.

Dataset S3 *Fragaria* × *ananassa cis* eQTL statistics for lead GWAS SNP AX123363242 andgenes within the 7A locus.

Dataset S4 Interproscan annotations for Medallion haplotype 1 genes within the 7A locus.

Dataset S5 BLASTp results of *FaECH* to all Medallion haplotype 1 protein sequences with average gene expression in eQTL and ripening contrast results from DESeq2.

Dataset S6 DESeq2 contrast results for 11 genes with AX123363242 eQTLs comparing green vs red condition in *Fragaria* × *ananassa* fruit ripening dataset.

Dataset S7 BLASTp result table from peach gene query sequences to the Medallion haplotype 1 protein annotations with average expression in the *Fragaria* × *ananassa* eQTL dataset and ripening upregulation results from fruit ripening dataset.

Dataset S8 *Fragaria vesca* GWAS results for γ-decalactone, γ-dodecalactone and γ-octalactone from both harvests and both models; SNPs were filtered by chromosome 3 and 7 and $P \leq$ 1e-06.

Dataset S9 *Fragaria vesca* gene expression for *FvECH*-like, *FvECH* and *FvFAD1* in fruit receptacle over different ripening stages adapted from (Härtl *et al.*, 2017).

Fig. S1 PCA of SNP data of all *Fragaria* × *ananassa* Genotypes.

Fig. S2 DeepLoc 2.0 protein subcellular localization results for *FaECH*.

Fig. S3 BLASTp and protein alignment of *FaECH* and Arabidopsis *CHY1*.

Fig. S4 Multiple sequence alignment of *FvECH*-like alleles for *Fragaria vesca* individuals with 2 alternative alleles at Fvb7 2509 419.

Fig. S5 Multiple sequence alignment of *FvECH*-like alleles for *Fragaria vesca* individuals with 0 alternative alleles at Fvb7 2509 419.

Fig. S6 GC-MS semi-quantitative results used for γ-decalactone quantification in *Fragaria vesca FvECH*-like overexpression transient assay.

Fig. S7 Alignment of cultivated strawberry and *Fragaria vesca* genomic regions for *FAD* and *ECH* genes.

Fig. S8 Multiple sequence alignment of *FvFAD1* alleles for *Fragaria. vesca* individuals with 2 alternative alleles at Fvb3 1103 870.

Fig. S9 Comparison of γ-lactone production among *Fragaria vesca* individuals with differing *FvFAD1* alleles.

Fig. S10 Multiple sequence alignment of *FvFAD1* alleles for *Fragaria vesca* individuals with 0 alternative alleles at Fvb3 1103 870.

Fig. S11 *FaECH-like* eQTL effect in *Fragaria* × *ananassa*.

Notes S1 Miscellaneous methods for genome assembly and annotation.

Notes S2 Volatile quantification method using SPME for *Fragaria* × *ananassa* and *Fragaria vesca* transient fruit assays.

Table S1 List of Genotypes used for GWAS and their inclusion in eQTL and PACE marker testing.

Table S2 List of ISOseq SRA accessions used in Medallion genome annotation.

Table S3 List of RNAseq SRA accessions used for *Fragaria* × *ananassa* eQTL analysis.

Table S4 List of RNAseq SRA accessions for *Fragaria* × *ananassa* Fruit Ripening Dataset.

Table S5 List of commercially synthesized DNA constructs.

Table S6 qPCR results from *Fragaria vesca* fruit transient assay replicates.

Table S7 Two-way ANOVA for lactones in *Fragaria* × *ananassa* by SNPs AX184908435 (3B) and AX123363242 (7A).

Table S8 Tukey HSD test for single and combined SNP groupings on lactone concentration in *Fragaria* × *ananassa* data.

Table S9 Genes involved in Peach lactone biosynthesis.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.